

Abstract

Systems and methods are provided for the generation of synthetic repeatable data. In an illustrative implementation, an exemplary data environment comprises at least one computing application for the management, manipulation, and generation of data. The computing application operates on a predefined set of rules to generate a data set, having N elements, using a deterministic generator function which when executed always produces the same set of data. The seed is used to position the generator to a particular point in its sequence. To regenerate any particular entry in the data set, the generator, using the seed as an input, is executed and the desired data is re-generated. The illustrative implementation also contemplates that the generation of data may be parallelizable as each element is generated independently of any others.